



**OCEAN FEATURES AND HUMPBACK WHALE FORAGING IN  
EAST AUSTRALIA: A LATENT STATE BOOSTING APPROACH**

Journal:	<i>Marine Mammal Science</i>
Manuscript ID:	MMSCI-3607
Manuscript Type:	Article
Date Submitted by the Author:	29-Nov-2012
Complete List of Authors:	Rankin, Robert; CSS-Dynamac Inc, ; NOAA National Ocean Service, NCCOS-CCMA-Biogeography Branch Walker, Emily; Pacific Whale Foundation, Giralt, Oriol; Associació Cetàcea, Maldini, Daniela; Okeanis, Veazey, Lindsay; Pacific Whale Foundation, Silva, Itana; State of Hawai'i, Division of Aquatic Resources Hutsel, Amanda; Pacific Whale Foundation, Macie, Anne; Pacific Whale Foundation, Kaufman, Greg; Pacific Whale Foundation,
Keywords:	Humpback whale, East Australia Current, Foraging, Hierarchical model, Boosting, Machine Learning, Regression Trees

SCHOLARONE™  
Manuscripts

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17

**OCEAN FEATURES AND HUMPBACK WHALE FORAGING IN EAST AUSTRALIA: A  
LATENT-STATE BOOSTING APPROACH**

ROBERT W RANKIN<sup>1,2</sup>, EMILY WALKER<sup>3</sup>, ORIOL GIRALT<sup>4</sup>, DANIELA MALDINI<sup>5</sup>, LINDSAY M VEAZEY<sup>3</sup>,  
ITANA SILVA<sup>6</sup>, AMANDA HUTSEL<sup>3</sup>, ANNIE MACIE<sup>3</sup>, GREG D KAUFMAN<sup>3</sup>

<sup>1</sup> CSS-Dynamac Inc., 10301 Democracy Lane, Suite 300, Fairfax, VA 22030, USA

<sup>2</sup> NOAA National Ocean Service, NCCOS-CCMA-Biogeography Branch, 1305 East-West Highway,  
SSMC-4, N/SCI-1, Silver Spring, MD 20910, USA

<sup>3</sup> Pacific Whale Foundation, 300 Maalaea Road, Wailuku, HI 96793, USA

<sup>4</sup> Associació Cetàcea, Marta Mata 3, Montgat, Barcelona, Spain 08390

<sup>5</sup> Okeanis, P.O. Box 853 Moss Landing, CA 95039, USA

<sup>6</sup> State of Hawaiï Division of Aquatic Resources, 130 Mahalani Street, Wailuku, Maui, HI 96793, USA

## 1 ABSTRACT

2 Meso-scale ocean features from the East Australia Current (EAC) have broad influence on the region's  
3 marine communities and top-predators. However, the relationship between the EAC and cetaceans has  
4 been described as a “complete mystery” (Suthers *et al.* 2011). We used novel predictive modelling  
5 techniques to investigate the relationship between meso-scale features and water-mass properties on the  
6 extralimital feeding of humpback whales (*Megaptera novaengliae*) in south east Australia using a long-  
7 term dataset from Twofold Bay, New South Wales. We used a hierarchical, latent-state model based on  
8 component-wise boosting to decompose variation into multiple non-linear environmental processes and  
9 observation error. Contrary to previous suggestions, feeding probability was associated with colder Sea  
10 Surface Temperature (SST), not warmer. Also, meso-scale ocean features, such as thermal fronts and  
11 cold-core eddies, were more important than SST to feeding probability. However, the most important  
12 variables were smooth interpolators of time and space. The results highlight the importance of structured  
13 modelling of time, space, observation, and environmental processes together when using machine  
14 learning techniques, such as boosted regression trees, in studies on the behavior and distribution of marine  
15 top-predators.

## 16 KEYWORDS

17 Humpback Whale; East Australia Current; Foraging; Hierarchical Model; Boosting; Machine Learning;  
18 Regression Trees;

19

## 1 INTRODUCTION

2 The East Australian Current (EAC) is a major driver of eastern Australian marine conditions and is  
3 changing under a warming climate (Ridgway 2007). The effects on biota are largely unknown, but could  
4 be considerable given the importance of the EAC to the lives of top marine predators (Hobday 2010,  
5 Young *et al.* 2011) and prey (Brandt 1983, Taylor *et al.* 2010). In a recent review, Suthers *et al.* (2011)  
6 described the association between the EAC and cetaceans as a “complete mystery”. We suspect that EAC  
7 ocean features, such as eddies, play a large role in the behavior, life-history, and distribution of shelf-  
8 associated and migratory populations of east Australian cetaceans. In this study, we explore one possible  
9 association between east Australia oceanographic features and the extralimital foraging behavior of a  
10 well-studied population of humpback whales (*Megaptera novaengliae*), the East Australia Group V Stock  
11 (EAGVS).

12 The ecological impact of the EAC is driven by complex interactions of water-mass properties (*e.g.*, Sea-  
13 Surface Temperature, SST) and ocean features (*e.g.*, eddies and fronts). For example, warm-core anti-  
14 cyclonic eddies serve as incubators of novel krill communities as they travel south from the EAC  
15 (Griffiths and Brandt 1983). In our study, we focus on how variation in SST and meso-scale ocean  
16 features is associated with extralimital feeding of EAGVS humpback whales. In particular, we revisit  
17 recent suggestions by Stamation *et al.* (2007) that high temperatures lead to more incidences of humpback  
18 whale feeding. Instead, we suggest that cooler temperatures and ocean features are more likely drivers of  
19 feeding behavior. In particular, we posit that ocean features deserve more attention because *i*) they have  
20 been important in understanding the movement and behavior of cetaceans in other contexts (Davis *et al.*  
21 2002, Ballance *et al.* 2006, Bost *et al.* 2009, Woodworth *et al.* 2011), *ii*) they have generally been under-  
22 represented in marine top-predators studies (Tremblay *et al.* 2009), and *iii*) they are less predictable under  
23 climate change. Understanding how cetaceans interact and behave with meso-scale features could change,  
24 counter or obfuscate the predicted effects of warming and the pole-ward shift of feeding areas.

1 Studying the association between ocean features and cetacean behavior is challenging, especially when  
2 generalizing over multiple years and variation in climate-ocean oscillations. Lacking long-term systematic  
3 data, we used the opportunistic observations from a long-term photo-identification dataset gathered from  
4 EAVGS humpback whales in Twofold Bay, New South Wales, south-east Australia. The dataset is part of  
5 a broader sight-resight study which includes other locations along the Australian coast and dates from  
6 1984 through the 2000's (Forestell *et al.* 2003, 2011). While foraging behavior was initially tangential to  
7 the studies' objectives, such opportunistic feeding observations have been influential in documenting the  
8 frequency of extralimital feeding among migratory humpback whales (Stockin and Burgess 2005,  
9 Stamation *et al.* 2007), which was previously considered rare. EAGVS humpback whales primarily feed  
10 in high-productivity sub-polar waters, but extralimital feeding may be an important supplement to  
11 subadults during migration (Swingle *et al.* 1993). Extralimital feeding opportunities may be more fleeting  
12 to the vagaries of oceanographic changes (Stamation *et al.* 2007) and therefore more sensitive to climate  
13 change. From a data analysis perspective, the extralimital aspect of the Twofold bay feeding may confer  
14 advantages over studies at primary feeding grounds, in that there may be greater contrast in behaviors  
15 (*e.g.*, whales seen not-feeding), as well as tighter connections between observations of feeding and local  
16 oceanographic phenomena.

17 The use of opportunistic and novel sources of information for marine mammal science is increasing  
18 (Tonachella *et al.* 2012, Best *et al.* 2012). However, when the source of data is not collected under a  
19 concerted sampling strategy, most of the assumptions of Frequentist statistical tests and subsequent p-  
20 values and confidence intervals are no longer reliable for inference (Greenland 2005, Rothman *et al.*  
21 2008). A similar situation occurs while behavior data is collected ancillary to whale watching or photo-  
22 identification expeditions: there are multiple objectives guiding investigators' decisions about which pods  
23 to follow or disengage. This mismatch in protocol and data can distort inferences about feeding and ocean  
24 features if biases arise in systematic ways. For example, juveniles humpbacks are more likely to feed  
25 extralimitally (Swingle *et al.* 1993) but are less likely to fluke up (Blackmer *et al.* 2000) and could be

1 unintentionally under-prioritized during photo-identification expeditions. Similarly, Doniol-Valcroze *et*  
2 *al.* (2012) warned that researchers will record more types of behaviors given more time following a pod,  
3 possibly leading to increased confirmations of feeding with increased time spent with a pod, and inflated  
4 “non-feeding” observations when certain pods are disengaged prematurely (Martin *et al.* 2005).

5 To handle these types of biases, we used a hierarchical latent-state model to jointly analyze the ecological  
6 process (ocean covariates of feeding probability) as distinct from observation error (differential ability to  
7 recognize a pod as feeding as the time spent with pods changes). Hierarchical models have become a  
8 preferred tool in the study of terrestrial and marine ecology (Clark and Gelfand 2006, Cressie *et al.* 2009,  
9 Moore and Barlow 2011, Halstead *et al.* 2012), and latent state models present a multiple regression  
10 alternative to non-parametric attempts at effort standardization, which can result in novel response  
11 variables with no apparent statistical distribution (Wiseman *et al.* 2011).

12 While most commonly employed in a Bayesian framework, we developed a latent-state model based on  
13 component-wise booting. Component-wise boosting combines the structured, conditional distribution of  
14 hierarchical models with the flexible fitting and regularization techniques of machine learning (Borisov *et*  
15 *al.* 2009, Schmid *et al.* 2010, Hutchinson *et al.* 2011, Hochachka *et al.* 2012, Mayr *et al.* 2012).

16 Component-wise boosting also facilitates the interweaving and regularization of a variety of predictor  
17 functions, including linear, smooth, and complex interactions, tailored to the ecological process under  
18 consideration. Foremost, we anticipated strong non-linearities in the relationship between feeding  
19 behavior and meso-scale features (Tew Kai *et al.* 2009, Dragon *et al.* 2010, Woodworth *et al.* 2011),  
20 which motivated the use of data-driven techniques such as conditional regression trees (Hothorn *et al.*  
21 2006). Secondly, we anticipated strong seasonal trends in feeding probability, arising due to seasonal  
22 variation in whales’ internal states (*e.g.*, health, migration status), which motivated the use of smooth  
23 temporal interpolators such as penalized splines. Third, we anticipated strong spatial trends, resulting  
24 from the haphazard spatial spread of focal follows, and more generally, as expected with most ecological

1 phenomena (Dormann *et al.* 2007, Beale *et al.* 2010, Robinson *et al.* 2011), necessitating smooth spatial  
2 interpolators such as bivariate splines (Kneib *et al.* 2009, Hothorn *et al.* 2010).

3 Component-wise boosting is one of the few techniques which can combine hierarchical modelling,  
4 flexible fits, and space-time interpolation. Ideally, observation biases and space-time variation should be  
5 addressed at the study-design phase through spatially-extensive sampling which controls for time of year  
6 and targets a broad contrast of oceanographic features. However, such studies are expensive, logistically  
7 challenging, and perhaps too late to the influence of climate change (Caro and Sherman 2011). In  
8 contrast, the Twofold Bay dataset is extensive in time. At a minimum, the evidence from this study can  
9 serve to identify important ocean phenomena and direct attention for systematic studies.

## 10 METHODS

### 11 *Study Area and Field Methods*

12 EAGVS humpback whales migrate annually along the eastern coast of Australia from their breeding areas  
13 along the Great Barrier Reef to their austral summer feeding areas in the Southern Ocean (Chaloupka and  
14 Osmond 1999). Twofold Bay is located along the southern extent of the coastal route in New South  
15 Wales. Photo-identification studies have been conducted at Twofold Bay since 1995 (with the exception  
16 of 2001) between mid-September to mid-November during the humpback whale southern migration. The  
17 protocols are detailed in earlier studies (Kaufman *et al.* 1993, Forestell *et al.* 2003, 2011).

18 Pods were intercepted haphazardly from research vessels, and opportunistically from researchers working  
19 on whale-watching vessels, generally within 3- 12 km from the coast. For the purposes of sight-resight  
20 objectives, pods were disengaged once all fluke IDs had been collected, or disengaged prematurely when  
21 there was little chance of a fluke-up event.

22 Survey days varied each year, from a minimum of 4 in 2006, a maximum of 50 in 2008, and a median of  
23 21 survey days per year. To match encounter data with remote-sensing data, only observations from 1998

1 onward were included in the feeding analysis. A total of  $n=980$  pods were encountered during this time,  
2 of which 28% were observed feeding.

### 3 *Response Variable*

4 Each observation was a pod encounter and its subsequent focal follow. Pods were assigned a behavior of  
5 “observed feeding” ( $Y=1$ ) or “not observed feeding” ( $Y=0$ ) based on visual cues. Individual identities  
6 within pods were not considered in this analysis, and repeat encounters of the same group or individuals  
7 were considered new, independent observations (*i.e.*, new data points). Pods were indexed to their Julian  
8 day-of-year, year, as well as their geospatial position. Pod encounters typically spanned some small  
9 distance, but because behaviors were haphazardly indexed to either time or position, we averaged the  
10 latitude and longitude (UTM) coordinates and assigned each pod encounter a mean pod location. We used  
11 these positions to extract geospatial predictor data. Dynamic predictors variables, such as SST and eddy  
12 covariates, were indexed to the nearest week of each pod encounter.

### 13 *Environmental Predictors*

14 The predictor variables are listed in Tab 1. The Marine Geospatial Ecology Toolset (MGET 0.8a30;  
15 Roberts *et al.* 2010) was used to obtain and process rasters for sea-surface height anomalies, eddies, SST  
16 and SST front locations. Two different SST products were used depending on the year in order to get the  
17 most cloud-free SST data for the region: the AVHRR Pathfinder (Casey *et al.* 2010) for the years 1998  
18 and 1999, and MODIS (Brown and Minnett 1999) for 2003 and from 2005 to 2010. Both data sets had a  
19 spatial resolution of  $\sim 4.6$  km/pixel with a temporal resolution of 8 days. Thermal fronts were derived  
20 through MGET using the Cayula-Cornillon single image edge detection algorithm on blended AVHRR  
21 and MODIS SST images (Cayula and Cornillon 1992, Ullman and Cornillon 2000). We selected a  
22 threshold of  $0.3125^\circ \text{C/km}$  to delineate fronts. The value was high compared to other cetacean studies,  
23 such as Gannier and Praca (2007) who used a  $0.1^\circ \text{C/km}$  gradient to detect fronts in the Mediterranean

1 Sea. The higher value was warranted given the large temperatures differences between the Coral Sea and  
2 Tasman Current.

3 Eddies were identified via the Okubo-Weiss eddy detection algorithm (Henson and Thomas 2008) using  
4 sea-surface height anomaly products (produced by Ssalto/Duacs and distributed by AVISO with support  
5 from CNES at <http://www.aviso.oceanobs.com/duacs/>). The algorithm identified cold-core cyclonic  
6 eddies from concave-down rings and warm-core anti-cyclonic eddies from concave-up rings. The R-based  
7 GIS library `rgeos` (Bivand and Rundel 2012, R Development Core Team 2012) was used to extract  
8 point values from rasters and calculate distance to fronts and eddies to Twofold Bay (UTM Zone 56  
9 South, WGS84 datum).

10 To investigate collinearity and redundancy of information among predictors, we calculated pairwise linear  
11 dependence with Pearson's  $r$  as well as pairwise mutual information with the coefficient of Reshef *et al.*  
12 (2011).

### 13 *Base Learners*

14 Predictor variables had smooth and flexible functional forms, as facilitated by the machine-learning  
15 technique of boosting. We used smooth penalized splines for time and space predictors (Julian day, time-  
16 with-pod, UTM) as well as pod-size. We used conditional regression tree (CRT; Hothorn *et al.* 2006) for  
17 oceanographic predictors (SST, distance to fronts and eddies, depth) to allow interactions and complex  
18 fits. One candidate CRT base-learner included distance to fronts, SST, distance to anti-cyclonic eddies  
19 and distance to cyclonic eddies, while another candidate base-learner included the same predictors but  
20 lacked eddies' polarity. This framework facilitated the ability for either base-learner to compete for  
21 inclusion in the boosting ensemble, while avoiding collinearity issues among all three eddy-based  
22 predictors in one base-learner. Random intercepts were modelled as ridge-penalized categorical effects,  
23 grouped by years.

### 24 *Hierarchical Framework*

1 In the latent state framework, the true state of a pod is denoted  $Z \in \{\text{feeding}, \text{not feeding}\}$ . The 0,1-  
 2 observations were considered  $n=980$  independent Bernoulli trials, whose parameter was the probability of  
 3 a pod engaged in feeding ( $\psi_{z=1}$ ), such that:

$$4 \quad Z_{0,1} | \psi \sim \text{Bern}(\psi_{z=1})$$

$$\psi_{z=1} = \left( \frac{e^{f_\psi(X)}}{1 + e^{f_\psi(X)}} \right)$$

5  $\psi_{z=1}$  has a logit link to the predictor function  $f_\psi(X)$ , hereafter referred to as the “feeding ensemble”, which  
 6 is composed of  $M$  base-learners with predictor variables  $X$ .  $X$  included oceanographic and space-time  
 7 predictors (Tab. 1).

8 The observed data  $Y$  are not necessarily equal to the state of feeding,  $Y \neq Z$ . Instead, there is some  
 9 intervening probability of detecting a pod feeding ( $p_{y=1|z}$ ) given that it is engaged in feeding. This type of  
 10 framework has seen widespread adoption in ecology, known in other contexts as zero-inflated binomial  
 11 models, site-occupancy models, or occupancy-detection models (Hall 2000, MacKenzie *et al.* 2003,  
 12 Royle and Kéry 2007, Hutchinson *et al.* 2011). The conditional distribution incorporates the observation  
 13 error:

$$14 \quad Y_{0,1} | p, Z \sim \begin{cases} \text{Bern}(p_{y=1}) & \text{if } z = 1 \\ 0 & \text{if } z = 0 \end{cases}$$

$$p_{y=1} = \left( \frac{e^{f_p(W)}}{1 + e^{f_p(W)}} \right)$$

15  $p$  also has a logit-link to a predictor function  $f_p(W)$ , hereafter referred to as the “observation ensemble”  
 16 composed of  $J$  base-learners with predictor variables  $W$ . As part of  $W$ , we included the time spent with a  
 17 pod, reasoning that more time spent with a pod should allow more opportunity to confirm feeding  
 18 behavior (Doniol-Valcroze *et al.* 2012). We also included a random-intercept grouped by year, to allow  
 19 intra-annual correlation in the ability to detect a pod feeding.

20 The log-likelihood ( $J$ ) of the latent-state model is:

$$1 \quad \ell(Y, f_\psi, f_p) = \sum_{i=1}^n \ell_i = \sum_{i=1}^n \left( \log(e^{y_i f_p + f_\psi} - y_i e^{f_p} + e^{f_p} - y_i + 1) - \log(e^{f_p} + 1) - \log(e^{f_\psi} + 1) \right) \quad (1)$$

2 In the above, the predictor functions  $f_\psi$  and  $f_p$  have the  $i$  subscript dropped for clarity but are indexed to  
3 each  $i^{\text{th}}$  observation.

#### 4 *Component-wise Boosting*

5 Boosting is the iterative assembly of weak base learners (*e.g.*, conditional regression trees) into a strong  
6 prediction functions,  $f_\psi$  and  $f_p$  (Schapire 1990). Our goal is to estimate the functional form of the  
7 prediction functions  $f$  as they mapped the predictor variables onto the expected response of whale  
8 foraging. In the case of a latent-state distribution with two ensembles,  $f_\psi$  and  $f_p$ , the algorithm combines  
9 gradient descent (Friedman 2001) and back-fitting (Hastie and Tibshirani 1990) during each boosting  
10 step: the former refers to minimizing the loss function, the negative log-likelihood (1), by fitting base  
11 learners to the negative partial derivatives  $-\frac{\partial \ell}{\partial f_\psi}$  and  $-\frac{\partial \ell}{\partial f_p}$ ; back-fitting refers to updating  $f_\psi$  conditional  
12 on the current value of  $f_p$  and vice-versa. The algorithm and partial derivatives are detailed in the  
13 Supplementary Material. The scientific literature is replete with detailed descriptions of the component-  
14 wise boosting algorithm as well as demonstrations through simulation and empirical study (Borisov *et al.*  
15 2009, Schmid *et al.* 2010, Hutchinson *et al.* 2011, Hochachka *et al.* 2012, Mayr *et al.* 2012). Hothorn *et*  
16 *al.* (2010) demonstrate the use of boosting to model a univariate species response with space, time and  
17 environmental base learners.

18 The advantages of boosting are *i*) automatic selection of importance predictors during model fitting, *ii*)  
19 regularization through “shrinkage” and early stopping of the ensemble building (Schmid *et al.* 2010), and  
20 *iii*) tuning of only a few “hyper-parameters” to ensure a sufficient model: the shrinkage parameters and  
21 the early stopping criteria. Furthermore, regression trees perform relatively well under collinearity of  
22 predictor variables (Hothorn *et al.* 2006, Dormann *et al.* 2012), which is expected among ocean  
23 phenomena.

1 According to Schmid and colleagues (2010), the shrinkage parameter has little importance on model  
2 performance, provided the value is sufficiently small ( $\nu=0.001$ ). They suggest choosing a constant, shared  
3 shrinkage rate and tune the stopping criteria for either ensemble,  $M_{final}$  and  $J_{final}$ , through cross-validation.  
4 We used 5-fold cross-validation (repeated 15 times) to grow ensembles until the cross-validation  
5 prediction statistics began to worsen. We used the most conservative (smallest ensemble) as judged by cv-  
6 AUC, cv-Brier score, or cv-loss function (1).

7 Analyses were run in the R statistical computing environment using code modified from the `mboost`  
8 library (Bühlmann and Hothorn 2007).

### 9 *Ecological Inference*

10 While machine-learning methods do not lend themselves to familiar hypothesis testing and p-values, the  
11 issue is moot in the face of collinearity among predictors when there is little guarantee of correct type I  
12 error rates for marginal tests of predictor significance, regardless of fitting technique (Dormann *et al.*  
13 2012). However, because each boosting iteration performs variable selection by fit, the frequency of  
14 selected predictors in the ensemble (weighted by the drop in the loss function) provides a natural means to  
15 rank variable importance.

16 We used a combination of marginal plots and two-way interaction plots to infer the functional form of  
17 ocean predictors and whale feeding. A 200-fold non-parametric bootstrap of observations was employed  
18 to estimate prediction intervals for marginal plots. The strength of interactions were ranked according to  
19 Elith *et al.* (2008).

## 20 RESULTS

21 Transformed and scaled predictors showed moderate to high collinearity according to the Maximum  
22 Information Coefficient and Pearson's  $r$ : cyclonic vs. anti-cyclonic eddies (0.79/0.45); Julian day vs. anti-  
23 cyclonic eddies (0.71/-0.30); SST and cyclonic eddies (0.66/-0.13); Julian day vs. cyclonic eddies (0.64/-

1 0.05); anti-cyclonic eddies and fronts (0.63/-0.14); SST vs. anti-cyclonic eddies (0.63/-0.38); Julian day  
2 vs. SST (0.59/0.66); cyclonic eddies and fronts (0.56/ -0.08); and depth versus anti-cyclonic eddies  
3 (0.53/0) and cyclonic eddies (0.52/-0.03); SST and fronts (0.47/ 0.19).

4 Pre-tuning suggested that both the feeding ensemble  $f_{\psi}$  and the observation ensemble  $f_p$  should have equal  
5 stopping criteria. The overall stopping criteria was 6342. The mean cross-validation Brier score was 0.13  
6 and mean cv-AUC was 0.89, categorized as “excellent discrimination” (Hosmer and Lemeshow 2000).

7 For the feeding ensemble,  $f_{\psi}$ , the rank of variable importance was: Julian year-of-day (32%), UTM's  
8 spatial spline (27%), distance to eddies (all types combined: 16%; eddies: 10%; cyclonic: 6%; anti-  
9 cyclonic: 0%), distance to fronts (13%), SST (7%), random intercept by year (3%), and depth (3%). For  
10 the detection ensemble, the random intercept by year was ranked highest (94%) then time-with-pod (6%).  
11 On average, there was a 0.63 probability of observing a pod in the feeding state (conditional on feeding).

12 Marginal plots agreed with the ranking of important variables, whereby higher ranking variables also had  
13 larger changes in the feeding propensity (on the logit scale) and narrower prediction intervals (Fig. 1). All  
14 variables showed mostly monotonic responses. Pods showed increased probability of feeding further  
15 offshore towards the north-east and in deeper waters (Fig. 2). Pods had lower probabilities of feeding with  
16 greater distances to eddies and fronts, and lower feeding probabilities with higher SST, consistent with  
17 expectations. Cold-core cyclonic eddies were selected in the feeding ensemble, but warm-core anti-  
18 cyclonic eddies were not. The most important variable, Julian day-of-year, showed a strong decline in the  
19 probability of feeding over the migration season into late October and November. Feeding detectability  
20 varied in a strongly non-linear fashion, with increasing ability to detect a pod feeding up until  
21 approximately 30 minutes, followed by a slight decline at longer intervals.

22 For the feeding ensemble,  $f_{\psi}$ , the rank (and scores) of two-way interactions were: SST and distance to  
23 fronts (11.7); distance to eddies and distance to fronts (9.0); SST and distance to eddies (7.15); distance to  
24 fronts and cyclonic eddies (5.4); SST and cyclonic eddies (2.9). Interactions were non-linear and reflected

1 similar patterns of the marginal responses, *e.g.*, with greater probability of feeding at closer proximity to  
2 fronts, eddies, and lower temperatures. The functional form of SST appeared to depend heavily on the  
3 distance to ocean features, showing flat response at close proximity to fronts, and a unimodal response at  
4 close proximity to eddies (Fig. 3).

## 5 DISCUSSION

### 6 *Biological Implications*

7 This study provides preliminary evidence of the influence of meso-scale features (eddies and fronts) and  
8 SST on the extralimital feeding of EAGVS humpback whales. Our results suggest that humpback whales  
9 tended to feed in closer proximity to fronts, cold-core eddies and areas of lower SST. There are two  
10 notable results of our study.

11 First, there is evidence that extralimital feeding among Australian humpbacks is associated with colder  
12 SST, not warmer, as suggested by Stamation *et al.* (2007). Second, cold-core cyclonic eddies appeared to  
13 be more important for whale feeding probability than warm-core anti-cyclonic eddies. The latter result is  
14 contrary to the amount of scientific attention that been focused on warm-core EAC eddies (Tranter *et al.*  
15 1980, Brandt 1983, Griffiths and Brandt 1983), while EAC cold-core eddies has received hardly any  
16 (Suthers *et al.* 2011). However, we can guess about a likely functional connection between extralimital  
17 foraging and cold-core eddies by analogy with other western-boundary currents, such as the California  
18 Current (Nishimoto and Washburn 2002) and the Kuroshio Current (Kimura *et al.* 2000), whereby cold-  
19 core eddies are important for promoting upwelling and aggregations of plankton and fish.

20 Unfortunately, our ability to make strong causal connections between ocean features and behavior is  
21 stymied by the collinearity and mutual information among predictors (Dormann *et al.* 2012, Maldini *et al.*  
22 2012). In particular, SST fronts are a measure of a variety of biological functioning, being a looser metric  
23 of all meso-scale features in this study, including cyclonic eddies, anti-cyclonic eddies, and large-scale  
24 currents such as the Tasman Front. It is not surprising that fronts would therefore have a greater apparent

1 importance in the feeding ensemble than eddies, as seen elsewhere (Hyrenbach *et al.* 2006), irrespective  
2 of biological functioning. Fronts and eddies, however, can have very different biological properties and  
3 functioning, the latter able to transport and incubate novel plankton communities over large distances  
4 from the EAC. On a regional scale, eddies can facilitate cross-shelf transport of nutrients and prey  
5 (Condie *et al.* 2011), while the local environment at front and eddy boundaries can support high prey  
6 densities through convergence and heightened productivity (Bost *et al.* 2009). Such patchy prey  
7 aggregations may lead to local prey densities that exceed the thresholds by which whales judge to engage  
8 in feeding behavior (Piatt and Methven 1992).

### 9 *Climate Change*

10 The distinction between water-mass properties and process-based ocean features (*e.g.*, meso-scale  
11 features; Tremblay *et al.* 2009) provides a fruitful framework to consider how whale behavior may be  
12 impacted by climate change. The east Australia shelf waters are suggested to have a high climate-velocity  
13 of approximately 40km/decade (Burrows *et al.* 2011), a rate higher than the global ocean average. It may  
14 be expected that the bottom-up signature of shifting prey distributions will negatively impact the foraging  
15 opportunities of whales. Humpback whales should only feed above certain prey densities (Piatt and  
16 Methven 1992), and given the suggested rarity of optimal conditions for feeding around Twofold Bay  
17 (Stamation *et al.* 2007), it may be that such extralimital zones are highly sensitive to climate change.

18 However, if ocean features, such as eddies and fronts, are more important for prey distribution and  
19 foraging opportunities, then the relationship between climate change, climate velocity, and prey response  
20 becomes more unpredictable. To our knowledge there is no definitive research on how the strengthening  
21 EAC will influence east Australian meso-scale features and their biological functioning. It may be that a  
22 stronger EAC could encourage greater opportunities for extralimital foraging if there is a more southerly  
23 extension, frequency, strength, or persistence of EAC derived features.

### 24 *Boosting and Ocean Features*

1 Despite the empirical importance of ocean features to the fine-scale movement patterns and behavior of  
2 top-predators, most large-scale, multi-year studies neglect such features in favor of water-mass properties  
3 (Tremblay *et al.* 2009). In order to develop better predictive models, marine analyses should explore and  
4 better incorporate ocean features. There are large practical difficulties to do so, and recent innovations  
5 such as component-wise boosting offer a particularly noteworthy framework.

6 Foremost, component-wise boosting techniques accommodate a wider repertoire of conditional  
7 distributions beyond “modelling of mean”, to allow deterministic modelling of variance, skew, quantiles,  
8 as well as mixture distributions (Rigby and Stasinopoulos 2005, Mayr *et al.* 2012). We think this will be  
9 especially pertinent to large-scale marine mammal observational studies of ocean features, given the  
10 lessons learned from tag-and-track studies of marine mammals, seabirds, and sea turtles: these studies  
11 commonly rely on study-specific non-parametric permutation or bootstrap routines (Tew Kai *et al.* 2009,  
12 Dragon *et al.* 2010, Woodworth *et al.* 2011). Often, the results of such studies suggest that ocean features  
13 alter the shape and spread of a distribution, rather than location (e.g., mean). Univariate bootstrapping and  
14 permutation analyses are useful for testing single hypotheses, but not for multiple regression when it is  
15 necessary, as in our study, to incorporate many predictors and control for space-time components and  
16 observation error (but see Dragon *et al.* 2010). Component-wise boosting presents a way to incorporate  
17 effects on the location, spread, and shape of distributions within a multiple regression framework (Mayr  
18 *et al.* 2012).

19 Secondly, component-wise boosting offers a suite of flexible base learners to approximate the non-  
20 linearities that can be expected when coercing non-continuous features (*e.g.*, ocean front: yes or no) into  
21 metrics for regression. For example, this study used one of the simplest operationalizations of ocean  
22 features: the “distance to” metric (Gannier and Praca 2007, Thorne *et al.* 2011). Aside from the potential  
23 for cross-shelf transport of nutrients and prey to Twofold Bay, there is no *a priori* reason to believe that  
24 “distance to a feature” is a general and valid metric of the biological connection among ocean features,  
25 prey and cetacean behavior. Rather, such distance measures are a modelling convenience, and in doing so

1 one should not assume linear relationships. Regressions trees and hinge functions that estimate sharp  
2 discontinuities can better approximate such non-linear effects of ocean features. These functions have  
3 already received considerable attention in terrestrial predictive models (Phillips and Dudík 2008) and are  
4 nearly synonymous with the development of boosting (Ridgeway 1999).

5 Third, the arbitrary and varied ways of measuring and operationalizing ocean features presents  
6 researchers with a high-dimensional problem of variable selection. The literature shows little consistency  
7 in ways to measure ocean features, ranging from current-based measures (Manderson *et al.* 2011), sea-  
8 surface height products (Dragon *et al.* 2010), water-mass clustering algorithms (Oliver *et al.* 2004),  
9 Lyapunov exponents (Tew Kai *et al.* 2009) and many more. In some instances researchers may be  
10 interested in a specific feature with a particular biological functioning, such as cold-core or warm-core  
11 eddies from the EAC. Without such *a priori* focus, there is no feature or measure which can be reliably  
12 assumed to be important for marine top-predators and included in modelling exercises. Rather, studies  
13 suggest that there are species- or guild-specific relationships (Ballance *et al.* 2006). With integrated  
14 variable selection and regularization, boosting techniques are particularly suited to the problem of high-  
15 dimensional data, even in extreme cases when there are more predictors than observations (Dudoit *et al.*  
16 2002, Schmid *et al.* 2010).

17 Fourth, the multitude of ways to operationalize ocean features leads to collinearity in predictive models.  
18 For example, Lyapunov Exponents are a general measure of many types of ocean features with different  
19 biological properties, such as eddies and small-scale coherent structures. Other types of measures, such as  
20 sea-surface height products, are much more focused to particular features and biological properties. Such  
21 variable-specificity and overlap of biological function reduces the ability for models to reliably estimate  
22 marginal significance of predictors for making inferences. No modelling paradigm can escape such  
23 issues, but boosted regression trees and conditional regressions trees may perform as well or better than  
24 most other methods under conditions of collinearity (Hothorn *et al.* 2006, Dormann *et al.* 2012).

1 However, the use of machine-learning techniques in the marine context needs to be balanced with  
2 important structural components. Ignoring components such as space and time and observational errors  
3 could result in over-fitting of oceanographic variables, as seen in other top-predators studies (Oppel *et al.*  
4 2011). In this study, the combined importance of all water-mass and ocean features in the model was just  
5 36%, in contrast to space and time components, such as Julian time of day and smooth spatial spline,  
6 which accounted for 60% of variable importance. Managing such spatio-temporal variation is  
7 increasingly recognized as crucial to ecological modelling (Hothorn *et al.* 2010, Robinson *et al.* 2011),  
8 and more rigorous hypothesis testing would ideally handle space and time effects at the planning stage  
9 through deliberate study-design. Such planning could also reduce collinearity among time and ocean  
10 variables, strengthening confidence in results. Nonetheless, through considered structuring of predictors  
11 and base learners, we suggest a means to utilize powerful fitting procedures, such as regression trees, in  
12 combination with structured components such as detection probabilities and space-time interpolators.

#### 13 ACKNOWLEDGEMENTS

14 Funding for this work was obtained from Pacific Whale Foundation members and supporters, as well as  
15 the Queensland Parks and Wildlife Service. Special thanks are due to Libby Eyre of Macquarie University  
16 and Ros and Gordon Butt of Cat Balou Cruises in Eden, New South Wales, for providing valuable photo  
17 images and logistical support. Field operations were conducted in compliance with all applicable laws and  
18 regulations, including annual permits for the approach and photo-identification of adults and juveniles.  
19 We would like regional authorities for their continued support, including the New South Wales (NSW)  
20 Animal Research Authority, the NSW Office of Environment and Heritage, the Queensland Department  
21 of Employment, the Queensland Environmental Protection Agency and Great Barrier Reef Marine Park  
22 Authority. Special thanks to Dr. Rebecca Hutchinson for providing information on the loss function and  
23 partial derivatives for occupancy-detection models.

24

## LITERATURE CITED

- 1  
2 Ballance, L. T., R. L. Pitman, and P. C. Fiedler. 2006. Oceanographic influences on seabirds and  
3 cetaceans of the eastern tropical Pacific: A review. *Progress In Oceanography* 69:360–390.
- 4 Beale, C. M., J. J. Lennon, J. M. Yearsley, M. J. Brewer, and D. A. Elston. 2010. Regression analysis of  
5 spatial data. *Ecology Letters* 13:246–264.
- 6 Best, B. D., P. N. Halpin, A. J. Read, *et al.* 2012. Online cetacean habitat modeling system for the US east  
7 coast and Gulf of Mexico. *Endangered Species Research* 18:1–15.
- 8 Bivand, R., and C. Rundel. 2012. rgeos: Interface to Geometry Engine - Open Source (GEOS). Version  
9 0.1.16. Available from <http://CRAN.R-project.org/package=rgeos>.
- 10 Blackmer, A. L., S. K. Anderson, and M. T. Weinrich. 2000. Temporal variability in features used to  
11 photo-identify humpback whales (*Megaptera novaeangliae*). *Marine Mammal Science* 16:338–  
12 354.
- 13 Borisov, A., G. Runger, E. Tuv, and N. Lurpionglukana-Strand. 2009. Zero-inflated boosted ensembles for  
14 rare event counts. *Advances in Intelligent Data Analysis VIII* 5572:225–236.
- 15 Bost, C. A., C. Cotté, F. Bailleul, *et al.* 2009. The importance of oceanographic fronts to marine birds and  
16 mammals of the southern oceans. *Journal of Marine Systems* 78:363–376.
- 17 Brandt, S. 1983. Pelagic squid associations with a warm-core eddy of the East Australian Current.  
18 *Australian Journal of Marine and Freshwater Research* 34:573–585.
- 19 Brown, O. B., and P. J. Minnett. 1999. MODIS infrared sea surface temperature algorithm theoretical  
20 basis document, Version 2.0, under contract number NAS5-31361. University of Miami, Miami,  
21 FL. 91 pp. Available from [http://modis.gsfc.nasa.gov/data/atbd/atbd\\_mod25.pdf](http://modis.gsfc.nasa.gov/data/atbd/atbd_mod25.pdf).
- 22 Bühlmann, P., and T. Hothorn. 2007. Boosting algorithms: Regularization, prediction and model fitting.  
23 *Statistical Science* 22:477–505.
- 24 Burrows, M. T., D. S. Schoeman, L. B. Buckley, *et al.* 2011. The pace of shifting climate in marine and  
25 terrestrial ecosystems. *Science* 334:652–655.

- 1 Caro, T., and P. W. Sherman. 2011. Endangered species and a threatened discipline: behavioural ecology.  
2 Trends in Ecology & Evolution 26:111–118.
- 3 Casey, K. S., T. B. Brandon, P. Cornillon, and R. Evans. 2010. The past, present, and future of the  
4 AVHRR pathfinder SST program. Pages 273–287 in V. Barale, J. F. R. Gower, and L.  
5 Alberotanza, eds. Oceanography from Space. Springer, Dordrecht, Netherlands.
- 6 Cayula, J.-F., and P. Cornillon. 1992. Edge detection algorithm for SST images. Journal of Atmospheric  
7 and Oceanic Technology 9:67–80.
- 8 Chaloupka, M., and M. Osmond. 1999. Spatial and seasonal distribution of humpback whales in the Great  
9 Barrier Reef. Pages 89–106 in J. A. Musick, ed. Life in the Slow Lane: Ecology and Conservation  
10 of Long-lived Marine Mammals. American Fisheries Society Symposium, Bethesda, MD.
- 11 Clark, J. S., and A. E. Gelfand. 2006. Hierarchical Modelling for the Environmental Sciences: Statistical  
12 Methods and Applications. Oxford University Press, Oxford, UK.
- 13 Condie, S. A., J. V. Mansbridge, and M. L. Cahill. 2011. Contrasting local retention and cross-shore  
14 transports of the East Australian Current and the Leeuwin Current and their relative influences on  
15 the life histories of small pelagic fishes. Deep Sea Research Part II: Topical Studies in  
16 Oceanography 58:606–615.
- 17 Cressie, N., C. A. Calder, J. S. Clark, J. M. V. Hoef, and C. K. Wikle. 2009. Accounting for uncertainty in  
18 ecological analysis: the strengths and limitations of hierarchical statistical modeling. Ecological  
19 Applications 19:553–570.
- 20 Davis, R. W., J. G. Ortega-Ortiz, C. A. Ribic, *et al.* 2002. Cetacean habitat in the northern oceanic Gulf of  
21 Mexico. Deep Sea Research Part I: Oceanographic Research Papers 49:121–142.
- 22 Doniol-Valcroze, T., V. Lesage, J. Giard, and R. Michaud. 2012. Challenges in marine mammal habitat  
23 modelling: evidence of multiple foraging habitats from the identification of feeding events in blue  
24 whales. Endangered Species Research 17:255–268.
- 25 Dormann, C. F., J. Elith, S. Bacher, *et al.* 2012. Collinearity: a review of methods to deal with it and a  
26 simulation study evaluating their performance. Ecography 35:1–20.

- 1 Dormann, C. F., J. M. McPherson, M. B. Araújo, *et al.* 2007. Methods to account for spatial  
2 autocorrelation in the analysis of species distributional data: a review. *Ecography* 30:609–628.
- 3 Dragon, A.-C., P. Monestiez, A. Bar-Hen, and C. Guinet. 2010. Linking foraging behaviour to physical  
4 oceanographic structures: Southern elephant seals and mesoscale eddies east of Kerguelen  
5 Islands. *Progress In Oceanography* 87:61–71.
- 6 Dudoit, S., J. Fridlyand, and T. P. Speed. 2002. Comparison of discrimination methods for the  
7 classification of tumors using gene expression data. *Journal of the American Statistical*  
8 *Association* 97:77–87.
- 9 Elith, J., J. R. Leathwick, and T. Hastie. 2008. A working guide to boosted regression trees. *The Journal*  
10 *of Animal Ecology* 77:802–813.
- 11 Forestell, P. H., G. Kaufman, and M. Chaloupka. 2003. Migratory characteristics of humpback whales  
12 (*Megaptera novaeangliae*) in Hervey Bay and the Whitsunday Islands, Queensland, Australia:  
13 1993-1999. Report to the Environmental Protection Agency, Queensland Parks and Wildlife  
14 Service, Australia (unpublished). 47 pp. Available from  
15 [http://www.pacificwhale.org/documentSetting/UserFiles/Forestell%20et%20al%202003%20QLD](http://www.pacificwhale.org/documentSetting/UserFiles/Forestell%20et%20al%202003%20QLD%2093-99%20report.pdf)  
16 [%2093-99%20report.pdf](http://www.pacificwhale.org/documentSetting/UserFiles/Forestell%20et%20al%202003%20QLD%2093-99%20report.pdf).
- 17 Forestell, P. H., G. D. Kaufman, and M. Chaloupka. 2011. Long-term trends in abundance of humpback  
18 whales in Hervey Bay, Australia. *Journal of Cetacean Research and Management (Special Issue)*  
19 3:237–241.
- 20 Friedman, J. H. 2001. Greedy function approximation: a gradient boosting machine. *Annals of Statistics*  
21 29:1189–1232.
- 22 Gannier, A., and E. Praca. 2007. SST fronts and the summer sperm whale distribution in the north-west  
23 Mediterranean Sea. *Journal of the Marine Biological Association of the United Kingdom* 8:187–  
24 193.
- 25 Greenland, S. 2005. Multiple-bias modelling for analysis of observational data. *Journal of the Royal*  
26 *Statistical Society. Series A (Statistics in Society)* 168:267–306.

- 1 Griffiths, F., and S. Brandt. 1983. Mesopelagic crustacea in and around a warm-core eddy in the Tasman  
2 Sea off eastern Australia. *Marine and Freshwater Research* 34:609–623.
- 3 Hall, D. B. 2000. Zero-inflated Poisson and binomial regression with random effects: a case study.  
4 *Biometrics* 56:1030–1039.
- 5 Halstead, B. J., G. D. Wylie, P. S. Coates, P. Valcarcel, and M. L. Casazza. 2012. “Exciting statistics”:  
6 the rapid development and promising future of hierarchical models for population ecology.  
7 *Animal Conservation* 15:133–135.
- 8 Hastie, T. J., and R. J. Tibshirani. 1990. *Generalized additive models*. Chapman & Hall/CRC, London,  
9 UK.
- 10 Henson, S. A., and A. C. Thomas. 2008. A census of oceanic anticyclonic eddies in the Gulf of Alaska.  
11 *Deep Sea Research Part I: Oceanographic Research Papers* 55:163–176.
- 12 Hobday, A. J. 2010. Ensemble analysis of the future distribution of large pelagic fishes off Australia.  
13 *Progress In Oceanography* 86:291–301.
- 14 Hochachka, W. M., D. Fink, R. A. Hutchinson, D. Sheldon, W. K. Wong, and S. Kelling. 2012. Data-  
15 intensive science applied to broad-scale citizen science. *Trends in Ecology & Evolution* 27:130–  
16 137.
- 17 Hosmer, D. W., and S. Lemeshow. 2000. *Applied logistic regression*. 2nd Edition. Wiley-Interscience  
18 Publication, New York, NY.
- 19 Hothorn, T., K. Hornik, and A. Zeileis. 2006. Unbiased recursive partitioning: a conditional inference  
20 framework. *Journal of Computational and Graphical Statistics* 15:651–674.
- 21 Hothorn, T., J. Müller, B. Schröder, T. Kneib, and R. Brandl. 2010. Decomposing environmental, spatial,  
22 and spatiotemporal components of species distributions. *Ecological Monographs* 81:329–347.
- 23 Hutchinson, R. A., L. P. Liu, and T. G. Dietterich. 2011. Incorporating boosted regression trees into  
24 ecological latent variable models. Pages 1343–1348 in W. Burgard and D. Roth, eds. *Proceedings*  
25 *of the Twenty-Fifth AAAI Conference on Artificial Intelligence*. Association for the

- 1 Advancement of Artificial Intelligence. Available from  
2 <http://www.aaai.org/ocs/index.php/AAAI/AAAI11/paper/viewFile/3711/4086>.
- 3 Hyrenbach, K. D., R. R. Veit, H. Weimerskirch, and G. L. Hunt. 2006. Seabird associations with  
4 mesoscale eddies: the subtropical Indian Ocean. *Marine Ecology Progress Series* 324:271–279.
- 5 IOC, IHO, and BODC. 2003. Centenary Edition of the GEBCO Digital Atlas. Intergovernmental  
6 Oceanographic Commission and the International Hydrographic Organization as part of the  
7 General Bathymetric Chart of the Oceans, British Oceanographic Data Centre, British  
8 Oceanographic Data Centre, Liverpool, UK.
- 9 Kaufman, G. D., B. A. Lagerquist, P. H. Forestell, and M. G. Osmond. 1993. Humpback Whales of  
10 Australia: A Catalogue of Individual Whales Identified by Fluke Photographs. 1st Edition.  
11 Queensland Department of Environment and Heritage, Brisbane, Queensland, Australia.
- 12 Kimura, S., H. Nakata, and Y. Okazaki. 2000. Biological production in meso-scale eddies caused by  
13 frontal disturbances of the Kuroshio Extension. *ICES Journal of Marine Science* 57:133–142.
- 14 Kneib, T., T. Hothorn, and G. Tutz. 2009. Variable selection and model choice in geosadditive regression  
15 models. *Biometrics* 65:626–634.
- 16 MacKenzie, D. I., J. D. Nichols, J. E. Hines, M. G. Knutson, and A. B. Franklin. 2003. Estimating site  
17 occupancy, colonization, and local extinction when a species is detected imperfectly. *Ecology*  
18 84:2200–2207.
- 19 Maldini, D., R. Eby, R. Scoles, M. P. Cotter, and R. W. Rankin. 2012. Haul-out behavior of California  
20 sea otters in a coastal wetland. *Northwestern Naturalist* 93:63–74.
- 21 Manderson, J., L. Palamara, J. Kohut, and M. Oliver. 2011. Ocean observatory data are useful for regional  
22 habitat modeling of species with different vertical habitat preferences. *Marine Ecology Progress*  
23 *Series* 438:1–17.
- 24 Martin, T. G., B. A. Wintle, J. R. Rhodes, *et al.* 2005. Zero tolerance ecology: improving ecological  
25 inference by modelling the source of zero observations. *Ecology Letters* 8:1235–1246.

- 1 Mayr, A., N. Fenske, B. Hofner, T. Kneib, and M. Schmid. 2012. Generalized additive models for  
2 location, scale and shape for high dimensional data—a flexible approach based on boosting.  
3 Journal of the Royal Statistical Society: Series C (Applied Statistics) 61:403–427.
- 4 Moore, J. E., and J. Barlow. 2011. Bayesian state-space model of fin whale abundance trends from a 1991  
5 – 2008 time series of line-transect surveys in the California Current. Journal of Applied Ecology  
6 48:1195–1205.
- 7 Nishimoto, M. M., and L. Washburn. 2002. Patterns of coastal eddy circulation and abundance of pelagic  
8 juvenile fish in the Santa Barbara Channel, California, USA. Marine Ecology Progress Series  
9 241:183–199.
- 10 Oliver, M. J., S. Glenn, J. T. Kohut, A. J. Irwin, O. M. Schofield, M. A. Moline, and W. P. Bissett. 2004.  
11 Bioinformatic approaches for objective detection of water masses on continental shelves. Journal  
12 of Geophysical Research 109:C07S04.
- 13 Opper, S., A. Meirinho, I. Ramirez, B. Gardner, A. F. O’Connell, P. I. Miller, and M. Louzao. 2011.  
14 Comparison of five modelling techniques to predict the spatial distribution and abundance of  
15 seabirds. Biological Conservation. *In press*.
- 16 Phillips, S. J., and M. Dudík. 2008. Modeling of species distributions with Maxent: new extensions and a  
17 comprehensive evaluation. Ecography 31:161–175.
- 18 Piatt, J. F., and D. A. Methven. 1992. Threshold foraging behavior of baleen whales. Marine Ecology  
19 Progress Series 84:205–210.
- 20 R Development Core Team. 2012. R: a language and environment for statistical computing. Version 2.15.  
21 R Foundation for Statistical Computing, Vienna, Austria. Available from [http://www.r-](http://www.r-project.org)  
22 [project.org](http://www.r-project.org).
- 23 Reshef, D. N., Y. A. Reshef, H. K. Finucane, *et al.* 2011. Detecting novel associations in large data sets.  
24 Science 334:1518–1524.
- 25 Ridgeway, G. 1999. The state of boosting. Computing Science and Statistics 31:172–181.

- 1 Ridgway, K. R. 2007. Long-term trend and decadal variability of the southward penetration of the East  
2 Australian Current. *Geophysical Research Letters* 34:5.
- 3 Rigby, R. A., and D. M. Stasinopoulos. 2005. Generalized additive models for location, scale and shape.  
4 *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 54:507–554.
- 5 Roberts, J. J., B. D. Best, D. C. Dunn, E. A. Treml, and P. N. Halpin. 2010. Marine Geospatial Ecology  
6 Tools: An integrated framework for ecological geoprocessing with ArcGIS, Python, R,  
7 MATLAB, and C++. *Environmental Modelling & Software* 25:1197 – 1207.
- 8 Robinson, L. M., J. Elith, A. J. Hobday, R. G. Pearson, B. E. Kendall, H. P. Possingham, and A. J.  
9 Richardson. 2011. Pushing the limits in marine species distribution modelling: lessons from the  
10 land present challenges and opportunities. *Global Ecology and Biogeography* 20:789–802.
- 11 Rothman, K. J., S. Greenland, and T. L. Lash. 2008. *Modern Epidemiology*. 3rd Edition. Lippincott  
12 Williams & Wilkins, Philadelphia, PA.
- 13 Royle, J. A., and M. Kéry. 2007. A Bayesian state-space formulation of dynamic occupancy models.  
14 *Ecology* 88:1813–1823.
- 15 Schapire, R. E. 1990. The strength of weak learnability. *Machine Learning* 5:197–227.
- 16 Schmid, M., S. Potapov, A. Pfahler, and T. Hothorn. 2010. Estimation and regularization techniques  
17 for regression models with multidimensional prediction functions. *Statistics and Computing*  
18 20:139–150.
- 19 Stamation, K. A., D. B. Croft, P. D. Shaughnessy, and K. A. Waples. 2007. Observations of humpback  
20 whales (*Megaptera novaeangliae*) feeding during their southward migration along the coast of  
21 southeastern New South Wales, Australia: identification of a possible supplemental feeding  
22 ground. *Aquatic Mammals* 33:165–174.
- 23 Stockin, K. A., and E. A. Burgess. 2005. Opportunistic feeding of an adult humpback whale (*Megaptera*  
24 *novaeangliae*) migrating along the coast of Southeastern Queensland, Australia. *Aquatic*  
25 *Mammals* 31:120–123.

- 1 Suthers, I. M., J. W. Young, M. E. Baird, *et al.* 2011. The strengthening East Australian Current, its  
2 eddies and biological effects — an introduction and overview. *Deep Sea Research Part II: Topical*  
3 *Studies in Oceanography* 58:538–546.
- 4 Swingle, W. M., S. G. Barco, T. D. Pitchford, W. A. McLellan, and D. A. Pabst. 1993. Appearance of  
5 juvenile humpback whales feeding in the nearshore waters of Virginia. *Marine Mammal Science*  
6 9:309–315.
- 7 Taylor, M. D., T. J. Mullaney, and I. M. Suthers. 2010. Mesoscale distribution of larval *Euphausia similis*  
8 in various water masses of the East Australian Current. *Deep Sea Research Part I: Oceanographic*  
9 *Research Papers* 57:1295–1303.
- 10 Tew Kai, E., V. Rossi, J. Sudre, *et al.* 2009. Top marine predators track Lagrangian coherent structures.  
11 *Proceedings of the National Academy of Sciences* 106:8245.
- 12 Thorne, L. H., A. J. Read, K. W. Urian, *et al.* 2011. The influence of dynamic oceanography on cetacean  
13 abundance and distribution in Onslow Bay, North Carolina. Presented at the Nineteenth Biennial  
14 Conference on the Biology of Marine Mammals, 27 November - 3 December 2011, Tampa, FL.  
15 Available from  
16 [http://www.marinemammalscience.org/index.php?option=com\\_content&view=article&id=390&I](http://www.marinemammalscience.org/index.php?option=com_content&view=article&id=390&Itemid=65&abstractID=1080)  
17 [temid=65&abstractID=1080](http://www.marinemammalscience.org/index.php?option=com_content&view=article&id=390&Itemid=65&abstractID=1080).
- 18 Tonachella, N., A. Nastasi, D. Maldini, G. D. Kaufman, and R. W. Rankin. 2012. Predicting trends in  
19 humpback whale (*Megaptera novaeangliae*) abundance using citizen-science. *Pacific*  
20 *Conservation Biology*. *In press*.
- 21 Tranter, D. J., R. R. Parker, and G. R. Cresswell. 1980. Are warm-core eddies unproductive? *Nature*  
22 284:540–542.
- 23 Tremblay, Y., S. Bertrand, R. W. Henry, M. A. Kappes, D. P. Costa, and S. A. Shaffer. 2009. Analytical  
24 approaches to investigating seabird–environment interactions: a review. *Marine Ecology Progress*  
25 *Series* 391:153–163.

- 1 Ullman, D. S., and P. C. Cornillon. 2000. Evaluation of front detection methods for satellite-derived SST  
2 data using in situ observations. *Journal of Atmospheric and Oceanic Technology* 17:1667–1675.
- 3 Wiseman, N., S. Parsons, K. A. Stockin, and C. S. Baker. 2011. Seasonal occurrence and distribution of  
4 Bryde’s whales in the Hauraki Gulf, New Zealand. *Marine Mammal Science* 27:E253–E267.
- 5 Woodworth, P. A., G. S. Schorr, R. W. Baird, *et al.* 2011. Eddies as offshore foraging grounds for melon-  
6 headed whales (*Peponocephala electra*). *Marine Mammal Science* 28:638–647.
- 7 Young, J. W., A. J. Hobday, R. A. Campbell, R. J. Kloser, P. I. Bonham, L. A. Clementson, and M. J.  
8 Lansdell. 2011. The biological oceanography of the East Australian Current and surrounding  
9 waters in relation to tuna and billfish catches off eastern Australia. *Deep Sea Research Part II:*  
10 *Topical Studies in Oceanography* 58:720–733.
- 11
- 12

1

## FIGURE CAPTIONS

2 *figure 1.* Marginal response of the prediction functions to model predictor variables: probability of  
3 feeding ( $f_{\psi}$ , logit scale) and probability of observation of feeding conditional on feeding state ( $f_p$ , logit  
4 scale). Grey intervals represent estimated +/- 1 SD bootstrapped prediction interval.

5 *figure 2.* Interpolated response of the probability of feeding ( $f_{\psi}$  logit scale) in relation to geographic space  
6 of the study area (penalized spatial spline) in Twofold Bay, New South Wales, Australia. Filled 'o'  
7 represent feeding observations and "x" represent observations of no-feeding.

8 *figure 3.* Marginal response of the logit-scale probability of feeding ( $f_{\psi}$ ) in relation to two-way  
9 interactions between predictors.

10

## SUPPLEMENTARY MATERIAL

$Y$  is an  $n$  sized vector of 0-1 observations.  $f_\psi$  and  $f_p$  are  $n$  sized vectors of the predicted feeding probability and conditional observation probability, respectively, on the logit scale. Component-wise boosting estimates the functional form of  $f_\psi$  and  $f_p$  while undergoing gradient descent to the global minimum of the empirical risk function: the negative log-likelihood of the latent state model. The log-likelihood is:

$$\ell(Y, f_\psi, f_p) = \sum_{i=1}^n \ell_i = \sum_{i=1}^n \left( \log(e^{y_i f_p + f_\psi} - y_i e^{f_p} + e^{f_p} - y_i + 1) - \log(e^{f_p} + 1) - \log(e^{f_\psi} + 1) \right) \quad (1)$$

The algorithm estimates the predictor functions through step-wise estimation and fitting of the negative gradient vectors,  $U$ , to base-learners,  $b$ . The negative gradient vectors are calculated from the negative partial derivatives of (1) with respect to either component  $f_\psi$  and  $f_p$ , respectively:

$$-\frac{\partial \ell}{\partial f_\psi} = \frac{e^{y f_p + f_\psi} + y e^{f_p + f_\psi} + y e^{f_\psi} - e^{f_p + f_\psi} - e^{f_\psi}}{e^{y f_p + 2 f_\psi} + e^{y f_p + f_\psi} - y e^{f_p + f_\psi} - y e^{f_\psi} + e^{f_p} (e^{f_\psi} - 1) + e^{f_\psi} - y + 1} \quad (2), \text{ and}$$

$$-\frac{\partial \ell}{\partial f_p} = \frac{e^{y f_p} (y e^{f_p + f_\psi} + y e^{f_\psi} - e^{f_p + f_\psi})}{e^{y f_p + f_p + f_\psi} + e^{y f_p + f_\psi} - y e^{2 f_p} - 2 y e^{f_p} + e^{2 f_p} + 2 e^{f_p} - y + 1} \quad (3), \text{ where } f \text{ and } y \text{ are indexed to each } i^{\text{th}}$$

observation.

The algorithm proceeds as follows, based on Schmid *et al.* (2010):

- 1)  $f_\psi^0$  and  $f_p^0$  are initialized as zero vectors  $0_n$ .
- 2) loop through  $m=1:M$  and  $j=1:J$ :
- 3) calculate the negative gradient vector  $U_\psi^m$  by evaluating the negative partial derivatives at  $f_\psi^{m-1}$  and  $f_p^{j-1}$ :

$$U_\psi^m = -\frac{\partial}{\partial f_\psi} \ell(Y, f_\psi^{m-1}, f_p^{j-1})$$

- 4) loop through candidate base learners,  $b$ , composed of predictor variables  $X_b$ , and perform a regression of  $U_\Psi^m$  upon  $b(X_b)$ , to yield fitted gradient vectors:  $\hat{U}_{\Psi,b}^m$
- 5) select the best fitting base learner  $b(X_b)$  by a goodness-of-fit criteria,  $\Sigma(U_\Psi^m - \hat{U}_{\Psi,b}^m)^2$ , and update the prediction function with the corresponding fitted gradient vector, multiplied by the shrinkage rate  $\nu$ :  $f_\Psi^m = f_\Psi^{m-1} + \nu \hat{U}_\Psi^m$
- 6) repeat steps 3 – 5 for the observation component,  $f_p^j$ , using its suite of base learners and predictor variables,  $W$ , and conditional on the updated value of  $f_\Psi$ .
- 7) Stop updating  $f_\Psi$  and  $f_p$  at  $M$  and  $J$  respectively.

The stopping criteria  $M$  and  $J$  and the shrinkage rate  $\nu$  are hyperparameters which control the size of the ensembles and the weight each base learner contributes to the overall prediction function.

Hyperparameters are tuned through k-fold cross-validation.

A variety of linear and non-linear base learners are available as part of the R packages `mboost` (Bühlmann and Hothorn 2007) and `party` (Hothorn *et al.* 2006). The R package `gamboostLSS` implements the component-wise boosting algorithm for an arbitrary distribution defined by the user with an appropriate empirical risk function (1) and negative partial derivatives (2, 3).

Table 1: Summary of predictor variables and base learners

Component	Predictor	Base Learner	Scale	Source	Processing
$Pr(\text{feeding})$ $f_w$	depth	conditional regression trees <sup>a,b</sup>	~92.5 m/pixel	General Bathymetric Chart of the Oceans (GEBCO) <sup>1</sup>	
	distance to anti-cyclonic eddies	conditional regression trees <sup>a</sup>	~37km/pixel 8-day	AVISO <sup>2</sup> sea surface height anomalies	Okubo-Weiss eddy detection algorithm <sup>3</sup> ; log-transformed
	distance to cyclonic eddies	conditional regression trees <sup>a</sup>	~37km/pixel 8-day	<i>ibid.</i>	<i>ibid.</i>
	distance to eddies	conditional regression trees <sup>b</sup>	~37km/pixel 8-day	<i>ibid.</i>	minimum of distance to cyclonic eddies or anticyclonic eddies
	distance to SST fronts	conditional regression trees <sup>a,b</sup>	~4.6 km/pixel, 8-days	AHRR Pathfinder <sup>4</sup> and MODIS <sup>5</sup> daytime SST	Cayula and Cornillon SST front detection algorithm <sup>6</sup> ; log-transformed
	SST	conditional regression trees <sup>a,b</sup>	~4.6km/pixel, 8-days	<i>ibid.</i>	
	Julian day-of-year	penalized spline <sup>c</sup>	day	<i>in situ</i>	
	intercept   year	random intercept by year (ridge-penalized categorical effects) <sub>d</sub>	annual	<i>in situ</i>	
	pod size UTMs	penalized spline <sup>c</sup> bivariate penalized spline <sup>f</sup>	counts	<i>in situ</i> <i>in situ</i>	
$Pr(\text{observed as feeding}   \text{feeding})$ $f_p$	time-with-pod	penalized spline <sup>g</sup>	decimal hours	<i>in situ</i>	
	intercept   year	random intercept by year (ridge-penalized categorical effects) <sub>h</sub>	annual	<i>in situ</i>	

a – f and g-h designate inclusion in competing base learners

<sup>1</sup> IOC *et al.* (2003)

<sup>2</sup> produced by Ssalto/Duacs and distributed by AVISO with support from CNES at <http://www.aviso.oceanobs.com/duacs>

<sup>3</sup> Henson and Thomas (2008)

<sup>4</sup> provided by GHRSSST and the US National Oceanographic Data Center, supported in part by a grant from the NOAA Climate Data Record (CDR) Program for satellites (Casey *et al.* 2010)

<sup>5</sup> Brown and Minnett (1999)

<sup>6</sup> Cayula and Cornillon (1992) and Ullman and Cornillon (2000)

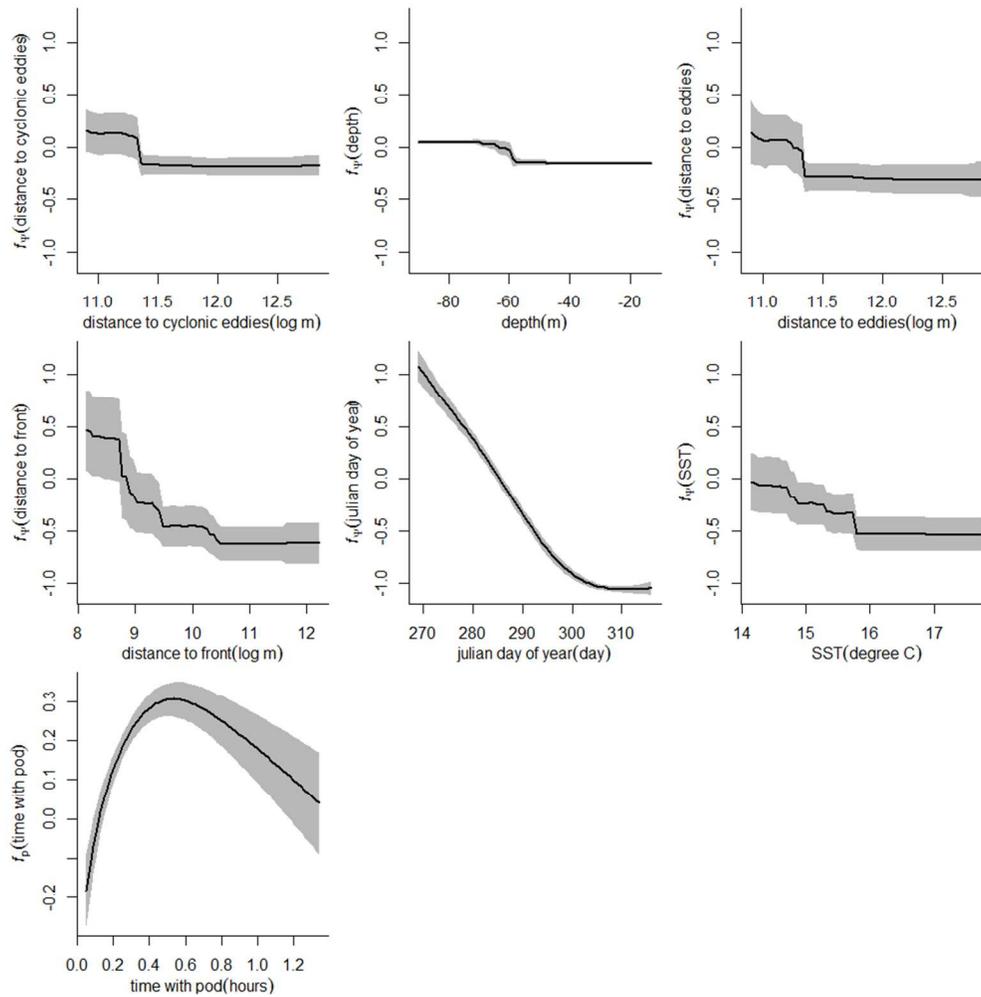
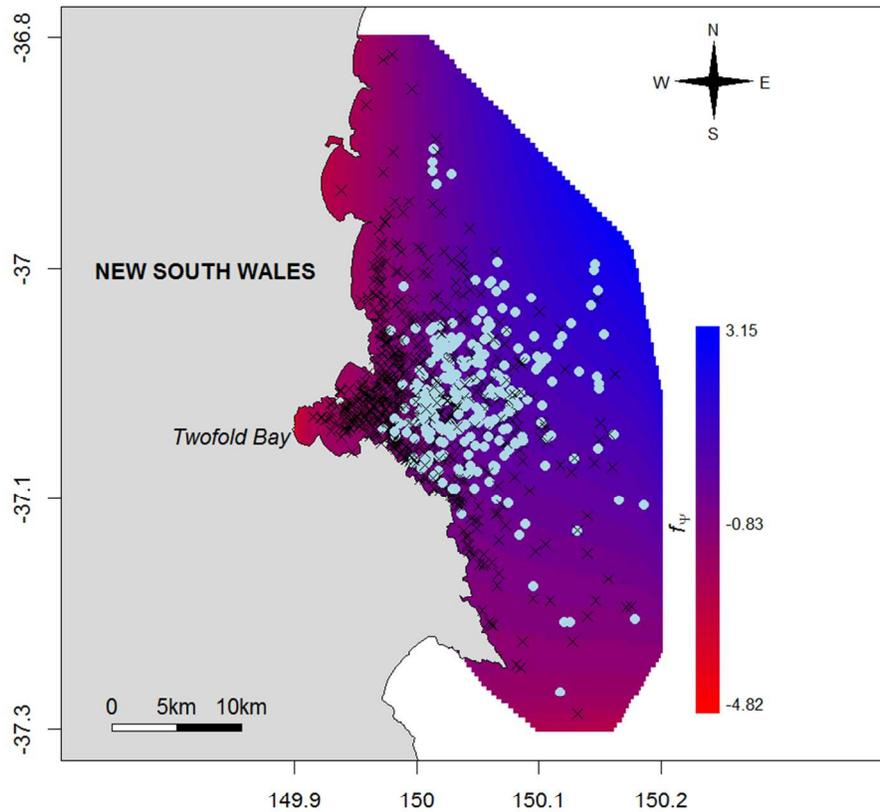
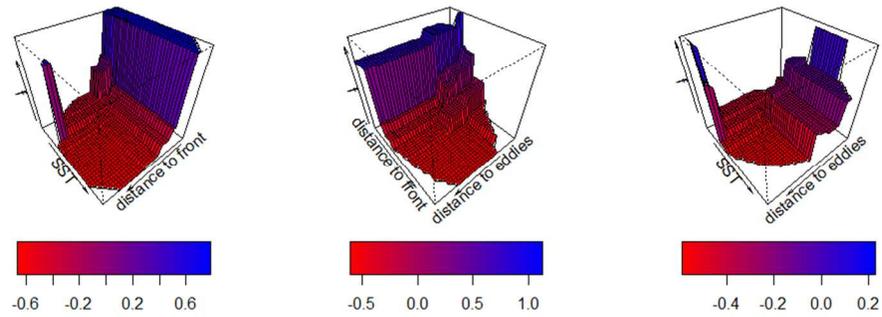


figure 1. Marginal response of the prediction functions to model predictor variables: probability of feeding ( $f_w$ , logit scale) and probability of observation of feeding conditional on feeding state ( $f_p$ , logit scale). Grey intervals represent estimated  $\pm 1$  SD bootstrapped prediction interval.  
352x352mm (72 x 72 DPI)



Interpolated response of the probability of feeding ( $f_p$  logit scale) in relation to geographic space of the study area (penalized spatial spline) in Twofold Bay, New South Wales, Australia. Filled 'o' represent feeding observations and "x" represent observations of no-feeding.  
352x352mm (72 x 72 DPI)



Marginal response of the logit-scale probability of feeding ( $f_w$ ) in relation to two-way interactions between predictors.

352x176mm (72 x 72 DPI)